2G1-5

Analyzing Optimal Marketing Strategies Over Customers' Networks

Qixin Liu, Yoshinobu Kawahara and Takashi Washio

The Institute of Scientific and Industrial Research (ISIR), Osaka University

In a costumers' network obtained, for example, from social networks, modeling influences among buyers is useful to estimate a buyer's decision to buy an item. This kind of influences is sometimes modeled with submodular functions and the problem of finding the marketing strategies to offer items to buyers that maximize the revenues of other buyers can be formulated as submodular maximization. In this paper, we address the problem of analyzing optimal and near-optimal solutions (strategies) to this problem and investigate numerically with real social network data.

1. Introduction

Social networking has been the most significant business development since recent years. Social networks, like Facebook, have allowed companies to collect information about the users and their relationships. Commonly, companies can make their marketing strategies over the social networks by means of advertising. However, research shows that the advertising-based business model has had only limited success on social networking sites. While a new business model other than the advertising-based business model which is known as viral marketing [Shapiro & Varian 99] is believed to have more effect and now is being studied by many researchers.

Imagine that Sony wants to promote its new digital camera. Sony can either advertise on Facebook and accept a very low click-through rate, or give away free cameras to several Facebook members (potentially at a lower cost than advertising) and generate a viral campaign. Research shows that this viral campaign is possible [Iyengar et al. 09]. The point is to find out the group of users who are more likely to be influenced by such a campaign and as well influence other friends to become buyers.

The fact that people are influenced by their friends is not new. We all know it at some level or the other. However, by understanding the social network of users, companies can better understand and influence consumers' behavior.

In order to monetize the marketing strategies, we have to consider the following two aspects: First, at what price should the products be given to users? There is a trade-off when offering products for free or giving discounts to a group of initial users. It decreases the revenue earned from the transaction while increasing the likelihood of a sale and the influence on future buyers. Second, in what sequence should the selling happen? Influence is usually not symmetric. We would like to find out the group of users to give free products to and the sequence to offer discounts which can cause further sales to occur earlier. Our goal is to investigate the marketing strategies that optimize a seller's revenue.

Such problems can be modeled as combinatorial problems but is very hard to find out the optimal solution. We notice that the influences is sometimes modeled with submodular functions and the problem of finding the marketing strategies to offer items to buyers that maximize the revenues of other buyers can also be formulated as submodular maximization problems [Hartline et al. 08]. To take advantage of such submodularity features, and by applying our cutting plane method [Kawahara et al. 09], we expect to gain the optimal and near-optimal solutions to the problem more efficiently and accurately. In this paper, we explore the optimal marketing strategies by means of submodular maximization and put forward some efficient algorithm to this problem.

We will give a detailed explanation of marketing strategies over social networks and show its hardness in Section 2.. Then we will show how to find the optimal marketing strategy by maximizing the expected revenue in the form of submodular function maximization. In Section 3., we will introduce our algorithm which is known as cutting plane method. In Section 4., the result of experiments with real social network data will be included. Finally we can reach a conclusion which shows our optimal and near-optimal solutions to this problem. And there are subjects left for further researches.

2. Marketing strategies and Submodular Maximization

As discussed in the introduction, a marketing strategy has the seller visiting buyers in a certain sequence and offering each buyer a price. We assume that the decision of a buyer whether to accept or reject an item is dependent on other buyers' behaviors (whether buy the items or not and at what price they buy the items).

Now we describe a general setting of the model. We assume that there is a set V of potential buyers. A buyer's decision to buy an item is dependent on other buyers who already own the item. For buyer $i \in V$, the value of the buyer for the good is defined by a set function $v_i : 2^V \to R^+$. The function models the influence that buyers have on future buyers. If a set $S \subseteq V \backslash \{i\}$ of buyers already owns the item, the value of buyer i, or the maximum willingness to pay of buyer i is a non-negative number $v_i(S)$. We assume that the seller doesn't know about the value functions but instead has distributional information F about them. $F(p)$

Contact: Qixin Liu, cherry@ar.sanken.osaka-u.ac.jp

is the cumulative distribution that the buyer's value is less than p. If we can find the optimal price p^* which can maximize the expected revenue $p\cdot(1-F(p))$ extracted from buyer i, then the optimal revenue will become $p^* \cdot (1 - F(p^*))$.

Here are some instantiations of the influence model that we use in the paper [Hartline et al. 08].

Uniform Additive Model

There weights w_{ij} for all $i, j \in V$. The value $v_i(S)$, for all $i \in V$ and $S \subseteq V\backslash\{i\}$, is drawn from the uniform distribution $[0, \sum_{j\in S\cup\{i\}} w_{ij}]$.

Symmetric Model

The valuation $v_i(S)$ is distributed according to a distribution F_k, where $k = |S|$.

Concave Graph Model

Each buyer $i \in V$ is associated with a non-negative, monotone, concave function $f_i : R^+ \to R^+$. The value $v_i(S)$ for all $i \in V$, $S \subseteq V\backslash\{i\}$, is equal to $f_i(\sum_{j\in S\cup\{i\}} w_{ij})$. Each weight w_{ij} is drawn independently from a distribution F_{ij}. The distributions $F_{i,S}$ can be derived from the distributions F_{ij} for all $j \in S$.

2.1 Optimal Price over Symmetric Settings

We now consider the optimal marketing strategy over symmetric settings and show the optimal price gained by a simple dynamic programming approach.

In the symmetric settings, we assume that the buyer values are drawn from one of $|V|$ distributions F_k. The offered price is only related with the number of buyers that have owned the item and the number of buyers who have not. Given that k people have accepted the item and t buyers remained under consideration, $p(k,t)$ is the offer price to the buyer under consideration, $R(k,t)$ is the maximum expected revenue that can be collected from the remaining buyers. The cumulative distribution $F_k(p)$ of the buyer's willingness to buy at a price less than p and its density function $f_k(S)$ also exist.

For a price p, if the buyer accepts, we can get the revenue of $p+R(k+1, t-1)$, and if he rejects, we can get the revenue of $R(k, t-1)$. The expected remaining revenue is

$$F_k(p) \cdot R(k, t-1) + (1 - F_k(p)) \cdot (R(k+1, t-1) + p)$$

By differentiating the expression with respect to p, we can get the optimal price.

$$f_k(p)(R(k, t-1) - R(k+1, t-1) - p) + 1 - F_k(p) = 0$$

By means of dynamic program we can get the optimal price over symmetric settings in polynomial time and thus get access to compute the expected revenue $R(k,t)$.

2.2 Hardness of Optimal Marketing Strategies

We now have got the optimal price over symmetric settings. We assume that the seller knows such buyers' values exactly. The only problem is to find the right sequence to offer item to buyers. If a set of buyers S have bought the item, the price to offer to the next buyer i is $v_i(S)$, which is known to the seller. Even so, the problem of finding the right sequence is NP-Hard. Here we consider the additive model where, $v_i(S) = \sum_{j\in S\cup\{i\}} w_{ij}$.

For the NP-Hard maximum feedback arc set problem [Berger & Shor 90], given an edge-weighted directed graph, we have to order the nodes of the graph to maximize the total weight of edges going in the backward direction in the ordering. Our optimal marketing strategies face the same NP-Hard problem. Let the nodes of the graph be the set of buyers, the edge weights be w_{ij}, the ordering to offer buyers be σ, and we offer the prices the same as the buyers' values. Then the revenue from such pricing is equal to the weight of the feedback arc set when the nodes in the graph are ordered in the reverse of σ. Thus finding the optimal marketing strategy is equivalent to computing the maximum feedback arc set problem and is NP-Hard [Hartline et al. 08].

2.3 Viral Campaign Marketing

Noticing the hardness problem in last section, we now try to simplify the model. As we have discussed in introduction, the viral campaign is considered as an effective marketing strategy. For viral campaign, the seller firstly gives away items for free to a selected group of buyers, then offers them the optimal price by a random sequence. The simplified model is relatively simple with only two extreme prices and random orderings.

We want to find a set A that maximizes the revenue $g(A)$ where $g(A)$ is the expected revenue when we give the item for free to the set A of buyers. We assume all the revenue functions R_i for $i \in V$ be non-negative, monotone and submodular, then the expected revenue function $g(A) = \sum_{i\in V\backslash A} R_i(A)$.$R_i(A)$ is a non-negative submodular set function where revenue function $R_i(A) = p\cdot(1-F(p))$. Consider the aggressive strategy, we use optimal pricing for the revenue function and $R_i(A) = v_i(S)(1 - F(v_i(S)))$. Thus $g(A) = \sum_{i\in V\backslash A} R_i(A) = \sum_{i\in V\backslash A} v_i(S)(1-F(v_i(S)))$. The proof of the non-negative submodularity of $g(A)$ is as follows:

It is obvious that $g(A)$ is non-negative and all we need to prove is that for any set $A \subseteq V$ and $C \subseteq V$, there exists

$$g(A) + g(C) \geq g(A \cap C) + g(A \cup C)$$

First, notice that throughout this paper we assume that R_i is monotone which means buyers only exert positive influence on each other, i.e. for all i and $A \subseteq B \subseteq V\backslash\{i\}$, $R_i(A) \leq R_i(B)$. Meanwhile R_i is submodular, i.e. for all i and $A \subseteq V, B \subseteq V\backslash A$, $R_i(A \cap B) + R_i(A \cup B) \leq R_i(A) + R_i(B)$, which implies that the marginal influence of one buyer on another decreases as the set of buyers who own the good increases. Noticing the monotonicity of R_i, for each $i \in (A\backslash C) \cup (C\backslash A)$:

$$\sum_{i\in A\backslash C} R_i(C) + \sum_{i\in C\backslash A} R_i(A) \geq \sum_{i\in A\backslash C} R_i(A\cap C) + \sum_{i\in C\backslash A} R_i(A\cap C) \quad (1)$$

Noticing the submodularity of R_i, for each $i \in V\backslash(A\cup C)$,

$$R_i(A) + R_i(C) \geq R_i(A \cup C) + R_i(A \cap C)$$

By summing the above two expression we have,

$$\sum_{i \in V \setminus (A \cup C)} R_i(A) + \sum_{i \in V \setminus (A \cup C)} R_i(C) \geq$$

$$\sum_{i \in V \setminus (A \cup C)} R_i(A \cup C) + \sum_{i \in V \setminus (A \cup C)} R_i(A \cap C) \tag{2}$$

By summing (1) and (2) we have

$$\sum_{i \in V \setminus A} R_i(A) + \sum_{i \in V \setminus C} R_i(C) \geq$$

$$\sum_{i \in V \setminus (A \cup C)} R_i(A \cup C) + \sum_{i \in V \setminus (A \cap C)} R_i(A \cap C)$$

In this paper, we address submodular function maximization under a cardinality constraint:

$$\max_{S \subseteq V} f(S) \qquad s.t. \, |S| \leq k,$$

where $V = \{1, ..., n\}$ and k is a positive integer with $k \leq n$. Notice that the function $g(A)$ is not monotone and we cannot use the simple greedy algorithm.

3. Algorithm

As stated above, designing optimal marketing strategies can be formulated as the problem of maximizing *non-monotone* submodular functions. For monotone submodular functions, it is known that the greedy method perform well [Nemhauser et al. 78]. However, since the revenue is non-monotone, the performance of the greedy method can be no longer assured. For this reason, here we apply a modified branch-and-bound method to this problem.

The branch-and-bound method is the today's state-of-the art method to solve a broad range of mathematical problems [Ibaraki 87]. In this method, we recursively generate sub-problems by fixing a part of variables in the parent problem (branching) and judging whether an optimal solution is included in the sub-problem and its further branched problems (bounding). For such judgement, we evaluate an upper bound of the subproblem ub^s and compare it with a lower bound of the original problem lb. That is, when ub^s is less than lb, it is assured that an optimal solution is not included in the subproblem and its further branched problems. Since submodular maximization can be formulated as a binary-integer linear program, *i.e.*, an optimal solution of submodular function is equivalent to the following:

$$\begin{aligned} \max \ & \eta \\ \text{s.t.} \ & \eta \leq g(S) + \textstyle\sum_{j \in V \setminus S} \rho_j(S) y_j \\ & \quad - \textstyle\sum_{j \in S} \rho_j(V \setminus j)(1 - y_j) \, ({}^\forall S \subseteq V), \\ & A_i^T \boldsymbol{y} \leq b_i (i = 1, \ldots, m), y_j \in \{0, 1\} (j \in V), \end{aligned}$$

where $\rho_j(S) := g(S \cup j) - g(S)$ $(S \subset V, \ j \in V \setminus S)$, we can design the way of calculating an upper bound for the bounding operation [Nemhauser & Wolsey 81].

1 Construct a BDD representing the size constraint.
2 Set $\mathcal{L} \leftarrow (\emptyset, V)$ and $lb \leftarrow f_g$, where f_g is the function value by the greedy algorithm.
3 **while** $\mathcal{L} \neq \emptyset$ **do**
4 Select a subproblem (A, V^s) from the top of the list $\mathcal{L}$ and remove it from $\mathcal{L}$ (let $\tau = |A|$).
5 Compute a greedy solution consisting of S plus additional elements from V^s. Let denote by lb^s its value and suppose the greedy algorithm generates the set $(S = S^\tau, \ldots, S^k)$.
6 Solve the linear programming:

$$\begin{aligned} \max \ & \eta \\ \text{s.t.} \ & \eta \leq f(S^t) + \sum_{j \in V \setminus S} \rho_j(S) y_j \\ & \quad - \sum_{j \in S^t} \rho_j(V \setminus j)(1 - y_j) (t = \tau, \ldots, k), \\ & A_i^T \leq b_i (i = 1, \ldots, m), y_j \in [0, 1] (j \in V^s), \end{aligned}$$

and set $ub^s = \max \ \eta$. If $lb^s > lb$, $lb \leftarrow lb^s$, $\bar{S} \leftarrow S$.
7 **if** $lb - \epsilon \geq ub^s$ **then**
8 Construct an exclusive cut for (A, V^s) and goto Step 4.
9 Construct the submodularity cut with respect to S^k and $lb - \epsilon$ and apply the conjunction operation between its BDD to the current BDD.
10 Append two sub-problems that correspond to the two edges from the current node in the BDD to $\mathcal{L}$.

Algorithm 1: The branch-and-bound algorithm with submodularity-cut applied for non-monotone submodular maximization.

Also, to accelerate the branch-and-bound method for our problem, we generate, at each iteration, a linear constraint $\boldsymbol{a}^T \boldsymbol{x} \leq b$ such that

$$g(S') \leq \gamma - \epsilon \ \text{ for all } \ S' \text{ s.t. } \boldsymbol{x}^{S'} \in \{\boldsymbol{x} \in \mathbb{B}^n | \boldsymbol{a}^T \boldsymbol{x} \leq b, \boldsymbol{x} \in P\},$$

where γ (=lb) is the current best solution value, P is the current feasible region and $\boldsymbol{x}^{S'}$ is the characteristic vector of subset S'.*1 Such constraint can be constructed using submodularity (*submodularity-cut*) (See [Kawahara et al. 09] for the detail). This could remove many feasible solutions in each branch-and-bound iteration.

In addition, in order to efficiently maintain and access solutions during the calculation, we implement the above framework with binary decision diagrams (BDDs). A BDD is a compact expression of a boolean function and enables us to calculate many algebraic operations on boolean functions efficiently in the space of BDDs [Akers 78, Knuth 09]. A feasible region or a linear constraint for set-function optimization can be represented using a BDD because both can regarded as a boolean function (ex. feasible→0 or infeasible→1). Thus, once both are represented using BDDs, we can conduct operations required in the branch-and-bound method in an efficient manner.

Based on the above, the pseudo-code of the algorithm we apply is described in Algorithm 1, where set (A, V^s) in list $\mathcal{L}$ represents the subproblem with ground set V^s and set

*1 The characteristic vector of $S (\subseteq V)$ is defined as $\boldsymbol{x}^S := \sum_{i \in S} \boldsymbol{e}_i (\in \mathbb{B}^n)$, where $\boldsymbol{e}_i$ is the i-th unit vector [Murota 03].

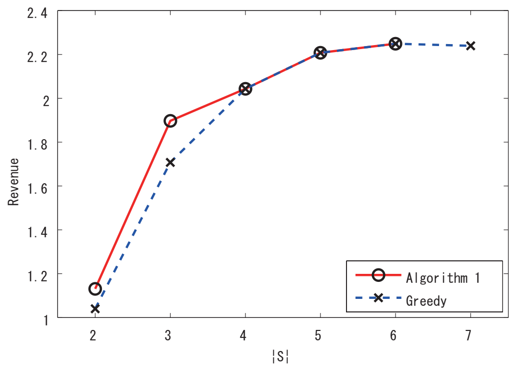

Figure 1: Revenus by Algorithm 1 and greedy method versus $|S|$ for a randomly generated subgraph with $|N| = 119$.

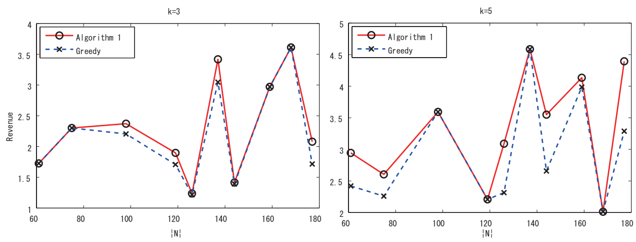

Figure 2: Revenus by Algorithm 1 and greedy method versus the sizes of randomly generated subgraphs for fixed k (3:left and 5:right).

A whose elements are assigned to 1. In each branch-and-bound loop, we solve the linear programming to evaluate an upper bound of a subproblem. In addition to an exact solution for non-monotone submodular maximization, Algorithm 1 can calculate an ϵ-optimal solution too (For an exact solution, ϵ can be set as 0).

4. Experiment and Evaluation

We applied Algorithm 1 to social network data and investigated empirically the discussed marketing strategies. To this end, we randomly generated moderate-size connected subgraphs using data from real social networks.*2

First, the graph in Figure 1 shows that the expected revenue grows with $|S|$, the number of users to offer free items to, and reaches the peak where the revenue is maximized, then goes down again with the growing size of $|S|$. Moreover, since the revenue function is not monotone, the greedy algorithm doesn't work well when selecting small $|S|$. But our algorithm gives better result in such cases.

Second, we fix the size of $|S|$ to 3 or 5 and compare these two algorithms with different sizes of nodes. The graphs in Figure 2 show that, for $|S| = 3$ or $|S| = 5$, our algorithm is better in getting exact values of the expected revenue than or equal to greedy algorithm.

5. Conclusion

In this paper, we worked on marketing strategies that optimize a seller's revenue considering the influence of social networks. We noticed that when the seller gives away items for free to a selected group of buyers, their friends are likely to be influenced by their purchasing behavior. And we formulated the optimal revenue the seller can get by choosing a group of users for such viral campaign as submodular maximization problem. Since the revenue function is not monotone, the greedy algorithm is no longer guaranteed to give the optimal solution. Thus we applied cutting plane method to this problem.

We conducted the experiment with data sets extracted from real social network data and make comparison between the performance of greedy algorithms and our algorithm. For selecting small size of nodes(3 or 5 nodes), our algorithm gives exactly the best answer while greedy algorithm cannot. When it comes to larger node size, these two algorithms generate almost the same result.

The problem is that it takes time to run the algorithm we proposed in this paper although it gives the exact answer. And for enumerating optimal marketing strategies, we still have to do experiments over larger data sets. Moreover, we will take into consideration the order and the optimal price to offer the items as well in our future work.

*2 We used the social network datasets available at `http://law.dsi.unimi.it/datasets.php` .

References

[Akers 78] S. B. Akers: Binary decision diagrams, IEEE Trans. on Computers **27**(6): 509-516, 1978.

[Berger & Shor 90] B. Berger and P. W. Shor: Approximation algorithms for the maximum acyclic subgraph problem, in Proc. of the first annual ACM-SIAM symp. on Discrete algorithms (SODA'90), pp.236-243, 1990.

[Hartline et al. 08] J. Hartline, V. Mirrokni and M. Sundararajan: Optimal marketing strategies over social networks, in Proc. of the 17th Int'l Conf. on World Wide Web (WWW'08), pp.189-198, 2008.

[Hassin & Rubinstein 94] R. Hassin and S. Rubinstein: Approximations for the maximum acyclic subgraph problem, Information Processing Letters **51**(3): 133-140, 1994.

[Ibaraki 87] Ibaraki, T.: Enumerative approaches to combinatorial optimization. In J. C. Baltzer and A. G. Basel (eds.), *Annals of Operations Research*, Vol.10-11, 1987.

[Iyengar et al. 09] R. Iyengar, S. Han and S. Gupta: Do Friends Influence Purchases in a Social Network? Retrieved from SSRN eLibrary, 2009.

[Jennifer et al. 06] J. A. Jennifer, D. W. Dahl and A. C. Morales: Consumer contamination: How consumers react to products touched by others, J. of Marketing **70**: 81-94, 2006.

[Kawahara et al. 09] Y. Kawahara, K. Nagano, K. Tsuda and J. Bilmes: Submodularity cuts and applications, Advances in Neural Information Processing Systems 22, pp.916-924, MIT Presss, 2009.

[Knuth 09] D. E. Knuth: Bitwise tricks & techniques: Binary decision diagrams. The Art of Computer Programming, Vol.4, 2009.

[Murota 03] K. Murota: *Discrete Convex Analysis*. Monographs on Discrete Math and Applications, SIAM, 2003.

[Nemhauser et al. 78] G. Nemhauser, L. Wolsey and M. Fisher: An analysis of the approximations for maximizing submodular set functions I, Mathematical Programming **14**(1): 265-294, 1978.

[Nemhauser & Wolsey 81] G. Nemhauser and L. Wolsey: Maximizing submodular set function: Formulation and analysis of algorithms, in P. Hansen (ed.), Studies on Graphs and Discrete Programming **11**, Annals of Discrete Mathematics, 1981.

[Shapiro & Varian 99] C. Shapiro and H. R. Varian: Information Rules: A Strategic Guide to the Network Economy. Harvard Business School Press, Cambridge, MA, 1999.

[Weimann 94] G. Weimann: The Influentials: People Who Influence People, Albany, NY: State University of New York Press, 1994.